数理解析研究所講究録
第816巻 1993年 116-126

Wavelets and their Applications

Mei Kobayashi and Masaharu Sakamoto
IBM Research, 5-19 Sanbancho, Chiyoda-ku Tokyo 102
e-mail: kobayash at trlvm.vnet.ibm.com
Workshop at Kyoto RIMS Nov.4-6, 1992

0. Introduction

Wavelets are receiving great attention for their ability to serve as a basis set for approximating functions and operators [D1],[D2],[Me1]. Although the technique shows promise in a variety of applications [Ch],[Co],[IEEE],[Me2],[MS],[R],[SJ], its advantages over conventional methods have not been clearly established, and further study is needed. For an excellent overview of the theory and an extensive bibliography, see [D3],[Me1]. This paper reviews works by others and describes our image and voice signal processing experiments using wavelets.

1. Image Processing

The wavelet concept was first introduced to image processing by Mallat [Ma1],[Ma2] whose algorithm is closely related to the Laplacian Pyramid Scheme on Burt and Adelson [BA]. The schemes compress data through a series of finer to coarser resolution spaces and allow efficient restoration of an image at a chosen resolution. Furthermore, the neighbourhood-based nature of the algorithms allows for fast local magnification of an image and edge detection during the compression process. Vetterli [V] and others have suggested that the ease of transition between resolution spaces is analogous to observing an object through a telescope and notes that the technique may be a useful tool for pattern matching or identification. Features and objects which consistently appear in images from consecutive resolution levels would play a key role in fast pattern or object matching and identification. Somewhat related to the multiresolution concept is the notion of fractals, which are characterized by self-similar patterns at different resolution levels. Arneodo et al. [A1],[A2],[A3] and Yamaguti [Y1],[Y2] have documented their work on fractal analysis using wavelets, noting in particular its application to the study of chaos and turbulence. Fractals have been used to generate various textures in computer graphics renditions so it is only natural that similar studies should be undertaken using wavelets. Froment and Mallat devised a *compact image coding algorithm that separates the edge from the texture information* [FM]. Chang and

Kuo [CK] analyzed and classified textures in images (e.g. canvas, straw cloth, wool, raffia) using a tree-structed wavelet. For a reference work which compares the short-time Fourier transform, Gabor expansions, wavelets and filter bank design for image processing applications, see [E].

Our image processing experiments using wavelets verified results from the literature and helped to develop a feel for the capabilities of the technique; no new results will be reported. Haar wavelets were selected since they correspond to the pixel-based representation of graphics screen systems. Copies of our color slides illustrating compression and edge detection for two-dimensional images are given in [Ko]. The slides were generated using image display software by Mr. Ioka [Io] of the Tokyo Scientific Center of IBM Japan on an IBM 6090 graphics terminal attached to a Canon Pixel Dio copier using a prototype converter box. Results from a second compression step are also given in [Ko] to show how the compression sequence proceeds.

Preliminary results from some three-dimensional image coding experiments which were shown at this workshop appear in [Ko]. The slides show the level set simulations from medical data for the skin, bone and combination skin/bone for the original and compressed data. During each step, the data is compressed by a factor of eight. Unfortunately, edge detection studies only yielded good results for two-dimensional images; the sparseness of the three-dimensional medical data and the different types of gradients between neighbouring tissues did not facilitate edge detection. The exception was the air-skin boundary, which could be easily recognized, and results from our studies were very good. We also experimented with two-dimensional video images where time was the third dimension, but was not able to produce meaningful results. Because of data overflow problems, we could not successfully extend level set method used to construct the three-dimensional medical images to visualize a series of frames from a simple black and white space walk scene from the movie *2001, A Space Odessey.* All calculations were performed using medical software provided by Mr. Miyazawa [My] of the Tokyo Scientific Center of IBM Japan on an IBM RS/6000 workstation which was connected to a Canon Pixel Dio machine via a prototype box.

Our experiments on image compression suggest that the wavelet technique, in the very least, requires substantial modification in order to compete with conventional methods. One promising application is the reconstruction of images using a series of successively finer edge detectors based on wavelet transform maxima [MZ],[Wil]. The technique might be used for efficient compression of video data or TV telephones since fast and reliable detection of edge movements within a given scene is vital for effective video signal processing. Another promising application is image analysis using wavelet transforms. The continuous, two-dimensional transform is used in [AMP] to *extract several descriptive parameters of images, ⋯ e.g. detection of position, orientation and visual contrast.* And an extension of the texture analysis work of Chang and Kuo [CK] to different types of objects may have important commercial and industrial applications.

2. Speech Signal Processing

The use of wavelets as an analysis and processing tool for sounds, music, speech and various acoustic signals has been documented at many workshops [Ch],[Co],[Me2],[R]. Scientists who have studied human speech processing using wavelet methods include Irino and Kawahara [I1],[I2],[IK1], [IK2],[Ka], Lee [L1],[L2], Tewfik and Sinha [ST1],[ST2],[TSJ], Sato [Sa1],[Sa2],[Sa3], Upton, Daimon, Isei, and Kunimatsu [UD],[UDIK] and Wickerhauser [Wic]. Irino and Kawahara's work focusses on the theory that wavelet methods are closely related to innate human sound processing mechanisms. Changes in the shape of the human vocal apparatus during speech are observed. Daniel Lee examined the compression and analysis of chirp signals. These signals are of interest because their rapid changes make them very difficult to compress, restore and analyze using conventional methods. In a comparison study of the conventional Gabor and wavelet transform, he found that *the Gabor transform cannot separate three component signals while the adaptive chirplet transform can* [L2]. Tewfik and Sinha developed an audio synthesis/coding method based on an optimization of the wavelet transform of a signal. In the references above, Sato discusses wavelets and his views on their signal processing capabilities. Upton, Daimon, Isei and Kunimatsu compare the digital filter and wavelet transform techniques, and apply their findings to the analysis of blasting noises and car door slams. Wickerhauser has studied the quality, efficiency and speed of acoustic signal compression using wavelet methods. The wavelets used by Irino and Lee are the same after re-arrangement of the constants. Upton and Wickerhauser do not give explicit formulae for the wavelets used in their studies.

Our first speech processing study examined acoustic compression rates and errors associated with Japanese human male speech using Haar and Daubechies wavelets of a uniform length in each run. Although fast algorithms appear in the literature, our calculations are based on a straightforward implementation of the wavelet methods described in the original works of Daubechies [D1],[D2]. The Haar wavelets we use are sets of characteristic functions having a support of 2^n points, where $n = 1, 2, 3, \ldots, 13$. Each compression step averages signals of pairs of neighbouring coefficients. The orthonormal wavelets with compact support described in [D1] were taken to have a support of

$3 \cdot 2^n$-points, where $n = 1, 2, 3, \ldots, 8$. Table 1 shows the number of coefficients needed to approximate a function sampled at $2^{14} = 16384$ evenly spaced points. We note that the 191 point wavelet minimizes the total number of coefficients to be stored. From the viewpoint of compression efficiency, larger sampling intervals favor longer wavelets, however longer wavelets require more edge coefficients for approximation near the endpoints. The two competing factors lead us to the minimum at the 191 point wavelet for a 16384 point interval. We caution the reader that different sampling intervals require an analogous calculation to determine the wavelet length for maximizing wavelet compression efficiency. We examined the sample phrase *Aoi Ume*, spoken by a Japanese male. Quantitative results from our experiments are given in Table 2. Qualitative results [KS1], e.g. our perceptions of the compressed sound, are consistent with our intuitive feelings; the Haar wavelet gave more corner-like, hard-edged sounds as the compression rate increased and the Daubechies a more spread-out, blurred, averaged sound.

In a second study, an edge and singularity detection wavelet proposed by Mallat and Hwang [MaH] was used to succesfully detect and analyze small bursts in hard consonant speech data. Accurate identification and analysis of these signals is important for speech signal processing and developing reliable voice input devices. They are often erased or can only be found through tedious work by Fourier based processing methods. We examined bursts in the isolated syllables: *aba, ada, aga, aka, apa, ata* (10 KHz sampling rate) as well as in words and phrases: *ta* in ki*ta*mi, *ta*kefumi, rikuzen*ta*kada, *ta*keo, ohmu*ta* and *ka* in *ka*zo, gushi*ka*wa, rikuzen*ta*kada, *ka*mo, mo*ka* (20 KHz sampling rate). Isolated syllable data is pure and simple, and can be used as a guide for study, however it tends to be contrived and different from sounds found embedded in natural speech. Color contour maps of the speech signals in time vs. wavelet resolution space were generated using the *Rubbersheet* subroutine in the IBM RS/6000 Data Visualizer Graphics Package. The location of the burst signals can clearly be seen [KS2] for the syllables *aga* and *aka*, but less so for for *aba, ada, apa and ata.* A finer sampling rate would have given a more pronounced mark for the latter four. When we doubled the sampling rate, the bursts for both *ka* and *ta* could clearly be seen in the complete words and phrases listed above. We are currently examining the analysis of burst signals using specialized wavelets.

In the future, we would like to examine speech analysis and pattern recognition capabilities of wavelets as well as possible relationships between fractals and wavelets. The authors would like to thank IBM Researchers Mr. Ioka and Mr. Miyazawa for use of

their graphics programs and helpful advice.

References

[A1] A. Arneodo et al., 182-196 in [Co].

[A2] A. Arneodo et al., 241-273 in [R].

[A3] A. Arneodo et al., 286-352, in [Me2].

[AMP] J-P. Antoine, R. Murenzi, B. Piete, 144-159 in [Me2].

[BA] P. Burt, E. Adelson, *IEEE Trans. Comm.*, 31(4), Apr. 1983, 532-540.

[CK] T. Chang, C.C.J. Kuo, *preprint, SIPI and Dept. EES at USC*, May 15, 1992.

[Ch] Chui, ed., *Wavelets*, Academic Press 1992.

[Co] J.M. Combes et al., eds., *Wavelets*, Springer 1989.

[D1] I. Daubechies, *Comm. Pure Appl. Math.*, 41,1988 , 909-996.

[D2] I. Daubechies, *IEEE Trans. Inf. Th.*, 36,1990 , 961-1005.

[D3] I. Daubechies, *Ten Lectures on Wavelets*, SIAM 1992.

[E] T. Ebrahimi, *Thesis*, Ecole Polytechnique Federale de Lausanne 1992.

[FM] J. Froment, S. Mallat 655-678 in [Ch].

[I1] T. Irino, *AVIRG Seminar on Wavelets* in Tokyo, Jan. 1992.

[I2] T. Irino, *Tech. Rep. IEICE*, Sp92-81, DSP92-66 ,Oct. 1992 59-66.

[Io] M. Ioka, DIS90G and DIS90C FORTRAN programs, IBM Research, Japan.

[IEEE] *IEEE Trans. Inf. Th.*, 38(2) March 1992 .

[IK1] T. Irino, H. Kawahara, *Proc. IEEE ICAASP 1992*, 15.3 Speech Coding.

[IK2] T. Irino, H. Kawahara, *Proc. Acoust. Soc. Japan 1992*, Spring(1-6-15), Fall(2-9-12)

[Ka] H. Kawahara, Chapter 4 in [SJ].

[Ko] Mei Kobayashi, IBM TRL/MSI Conf. report, July 17, 1992.

[KS1] M. Kobayashi, M. Sakamoto, *demo*, IBM Research, Japan, May 1992.

[KS2] M. Kobayashi, M. Sakamoto, *presentation*, Wkshp., Kyoto U. RIMS, Nov. 6, 1992.

[L1] D. Lee, *Wavelet Theory (preprint)*, HP Labs Japan, May 1992.

[L2] D. Lee, *Wavelet Course Notes*, HP Labs Japan, May 1992.

[Ma1] S. Mallat, *IEEE Trans. PAM*, 11(7) July 1989, 674-693.

[Ma2] S. Mallat, *IEEE Trans. ASSP*, 37(12) Dec. 1989, 2091-2110.

[MaH] S. Mallat, W-L. Hwang, *IEEE Trans. Inf. Th.*, 38(2) March 1992, 617-643.

[Me1] Y. Meyer, Ondelettes et Operateurs, Hermann 1990.

[Me2] Y. Meyer, ed. *Wavelets and Applications*, Masson 1992.

[My] T. Miyazawa, DISPIMG and VOXELVIEWER programs, IBM Research, Japan 1991 .

[MS] *Mathematical Sciences* (in Japanese) No. 354, Dec 1992.

[**R**] M.B. Ruskai et al, eds., *Wavelets*, Jones and Bartlett 1992.

[**Sa1**] M. Sato, *AVIRG Seminar on Wavelets* in Tokyo, Jan. 1992.

[**Sa2**] M. Sato, *ATR AVP Lab Tech. Pub.* Vol. 27, 1991, 72-75.

[**Sa3**] M. Sato, *Acoustical Soc. Japan* 47(6), 1991, 405-423.

[**SJ**] *SIAM Japan Seminar on Wavelets* (textbook).

[**ST1**] D. Sinha, A. Tewfik, *IEEE Int'l Conf. ASSP*, San Francisco, March 1992 .

[**ST2**] D. Sinha, A. Tewfik, *IEEE Sig. Proc.*, June 1993 , submitted.

[**TSJ**] A. Tewfik, D. Sinha, P. Jorgensen, 747-765 in [IEEE].

[**UD**] R. Upton, S. Daimon, *Jidousha Gijutsu Kai*, 921, 1992-5 (921013) 49-52.

[**UDIK**] Upton et al.,*Proc. Acoust. Soc. Japan 1992*, Fall(1-2-14) and (1-2-15).

[**V**] M. Vetterli, *IEEE Trans Sig Proc.*, 40(9) Sept. 1992 , 2207-2232.

[**Wic**] V. Wickerhauser, 679-700 in [Ch].

[**Wil**] R. Wilson et al., IEEE Trans. Info. Theory 38(2) 1992 , 674-690.

[**Y1**] M. Yamaguti, Chapter 1 in [SJ].

[**Y2**] M. Yamaguti, 5-10 in [MS].

Remark 1: A wavelet news net has been established in the hope of facilitiating the exchange of information on wavelets. Notices of conferences, papers, journals, general announcements as well as questions and answers are accepted. To access the archive site, anonymous ftp to maxwell.math.scarolina.edu (129.252.12.3), directory /pub/wavelet/archive. Gopher server: bigcheese.math.scarolina.edu. The number of subscribers is over 2300.

To submit, e-mail: wavelet at math.scarolina.edu with subject *submit.*

To subscribe, e-mail: wavelet at math.scarolina.edu with subject *subscribe.*

Remark 2: An increasing number of researchers have been placing postscript or dvi versions of their papers on-line, available by anonymous ftp. If you cannot interactively ftp the archive site, don't despair. The papers can be retrieved using a special node passing through Princeton. A sample file is given in Appendix 1. If you, yourself, plan to make your work available through anonymous ftp, it is strongly advised that long papers be cut into smaller chapter or section files. Large files may be purged by an intermediate system maintenance manager or automatically split into small pieces. I have experienced both; in the former case, I received a note of reprimand from an automatic purge mechanism, in the latter, the file had to be re-assembled by hand after retrieval, a tedious and time-consuming chore.

Remark 3: An Approximation Theory News Net (AT-NET) has been established in the hope of facilitiating the exchange of information on approximation theory. Notices of conferences, papers, journals, general announcements as well as questions and answers are accepted.

To submit, e-mail: maprx99 at technion.technion.ac.il

For questions, comments, e-mail: at-net at technion.technion.ac.il

Appendix 1:

To: bitftp pucc.princeton.edu at cunyvm.cuny.edu

ftp ltssun3.epfl.ch (address of node to be accessed)

anonymous

mei

dir (list contents of home directory)

cd pub/THESIS

dir

binary (postscript files are in binary form)

get thesis.ps.Z

quit

Table 1: $(3 \cdot 2^n - 1)-$ **point wavelet on 16384 points**

n	$3 \cdot 2^n - 1$	points per unit length	number of wavelet coeff	number of edge coeff	total no. of coeff
1	5	2	8190	8	8198
2	11	4	4094	16	4110
3	23	8	2046	32	2078
4	47	16	1022	64	1086
5	95	32	510	128	638
6	191	64	254	256	510
7	383	128	126	512	638
8	766^+	256	62	1024	1086

+ the 767 point wavelet equals zero at the 767^{th} point in doubleprecision arithmetic

Aoi Ume (Male Voice)

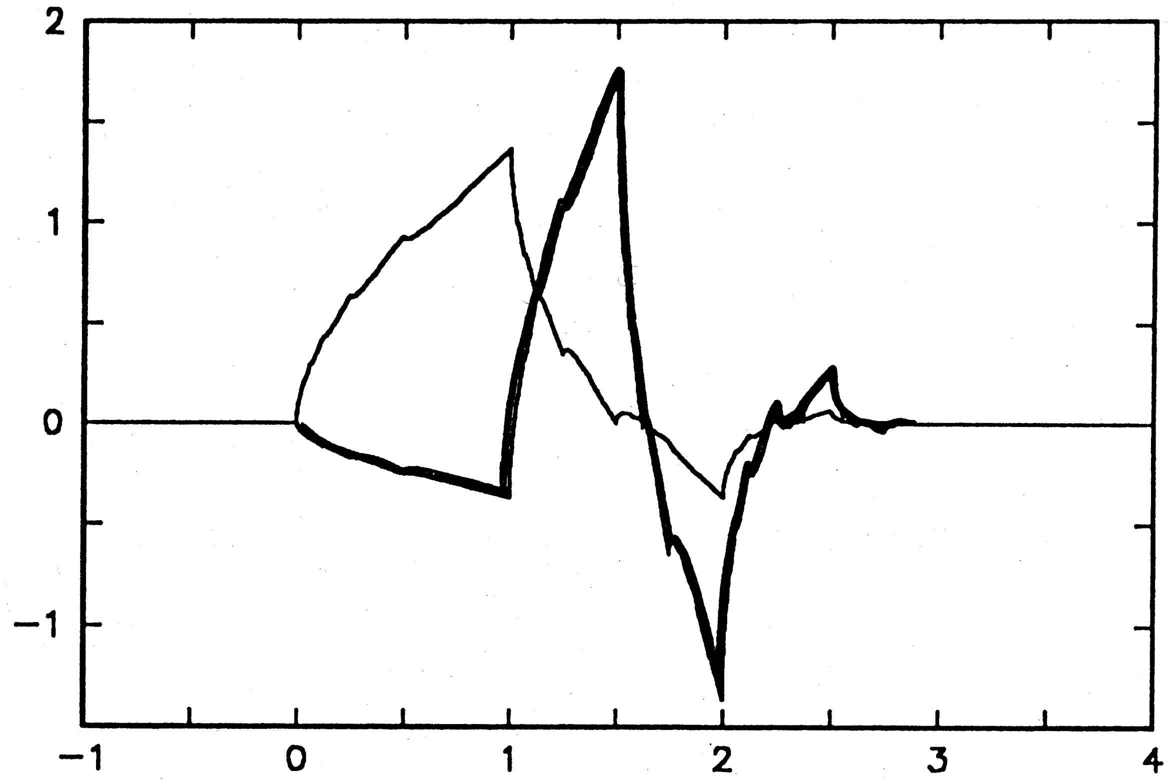
Table 2a: Haar Wavelet

coeff	wavelet	$L_\infty-$ error	x_∞	e_{total}	$e_{average}$
8192	2	24283	39	3887705	237
4096	4	34407	42	6044161	369
2048	8	32689	42	8927025	545
1024	16	30876	40	11714131	715
512	32	30778	40	12942017	790
256	64	29988	48	13405397	818
128	128	31156	40	13571981	828
64	256	32199	40	13706719	837
32	512	31731	40	13834123	844
16	1024	31077	40	14158927	864
8	2048	30694	40	14439561	881
4	4096	30501	40	14594229	891
2	8192	30404	40	14364489	877

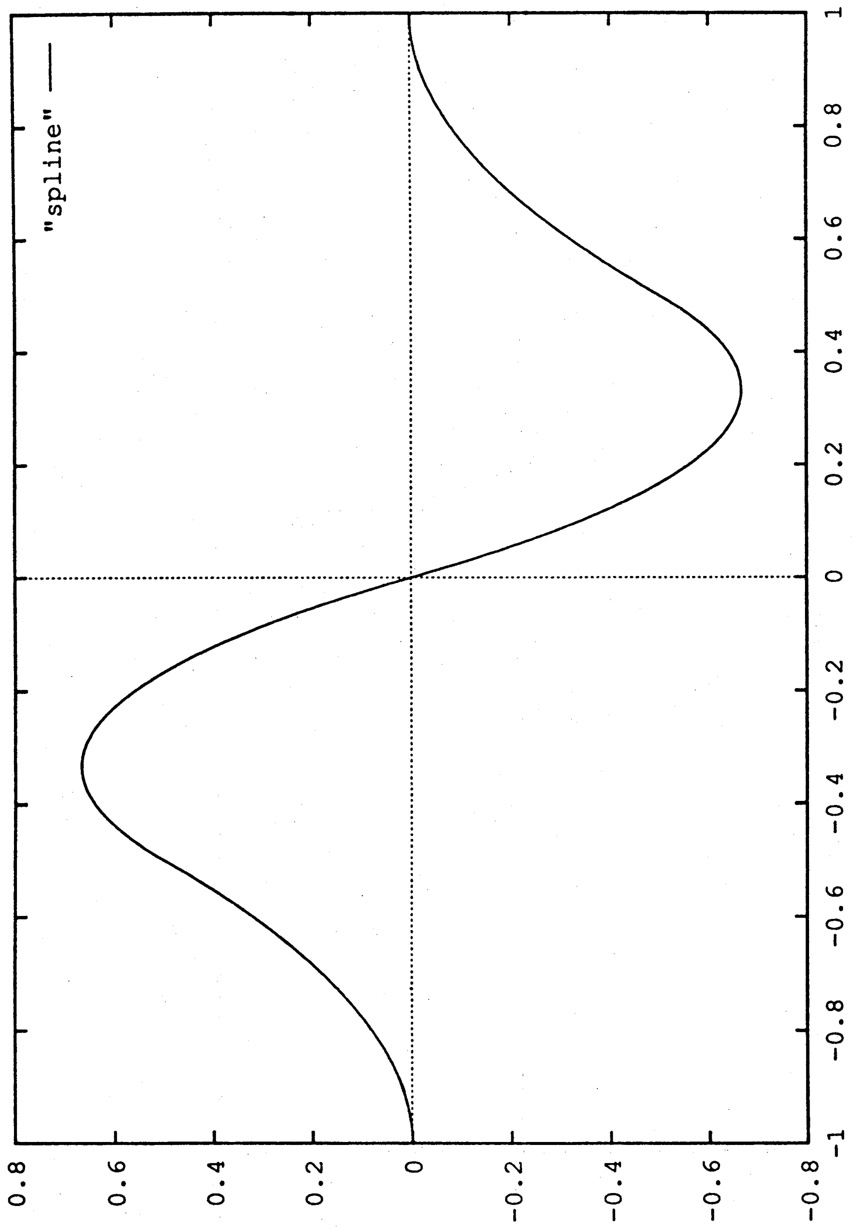
Table 2b: Daubechies Wavelet

coeff	wavelet	$L_\infty-$ error	x_∞	e_{total}	$e_{average}$
8198	5	30761	40	4119243	251
4110	11	32347	41	5588211	341
2078	23	28715	41	8277683	505
1086	47	32633	39	11150852	681
638	95	12325	6658	12112555	739
510	191	12103	6658	12403087	757
638	383	12042	6727	12427640	759
1086	+766	12052	6658	12433501	759

+ the 767 point wavelet equals zero at the 767^{th} point in doubleprecision arithmetic



Daubechies Wavelet and Scaling Functions



Mallat and Hwang Spline Wavelet